

EUROPEAN PATENT OFFICE

Patent Abstracts of Japan

PUBLICATION NUMBER : 11053392
PUBLICATION DATE : 26-02-99

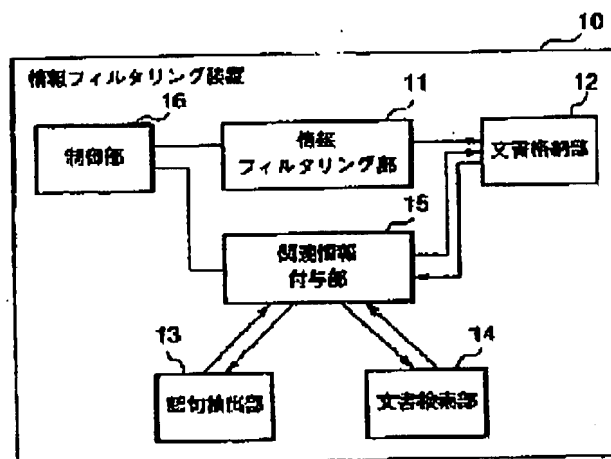
APPLICATION DATE : 08-08-97
APPLICATION NUMBER : 09214563

APPLICANT : TOSHIBA CORP;

INVENTOR : SUMITA KAZUO;

INT.CL. : G06F 17/30 G06F 12/00

TITLE : INFORMATION FILTERING DEVICE
AND RELATED INFORMATION
PRESENTATION METHOD APPLIED
TO THE DEVICE



ABSTRACT : PROBLEM TO BE SOLVED: To provide an information filtering device that makes it possible to easily access to a related information from information selected by a retrieval condition registered in advance.

SOLUTION: A related information assignment part 15 delivers a document which is selected by an information filtering part 11 and stored in a document storage part 12 to a words and phrases extraction part 13 and makes words and phrases in a specific category extracted from within this document. Next, the related information assignment part 15 delivers words and phrases which the words and phrases extraction part 13 extracts to a document retrieval part 14 and makes a retrieval of a related document on the basis of these extracted words and phrases performed. Then, the related information assignment part 15 adds related information including a storage position of the related document which this document retrieval part 14 retrieves to a selected document stored in the document storage part 12. Thus, a load of a user spent for a retrieval work of the related information for information presented as a filtering result can be largely reduced.

COPYRIGHT: (C)1999.JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-53392

(43) 公開日 平成11年(1999) 2月26日

(51) Int.Cl.ⁿ
G 0 6 F 17/30
12/00 5 4 7

F I
C 0 6 F 15/403 3 0 0 C
12/00 5 4 7 H
15/40 3 7 0 A
15/401 3 2 0 C

審査請求 未請求 請求項の数10 O L (全 9 頁)

(21) 出願番号 特願平9-214563

(22) 出願日 平成9年(1997) 8月8日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 住田 一男

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

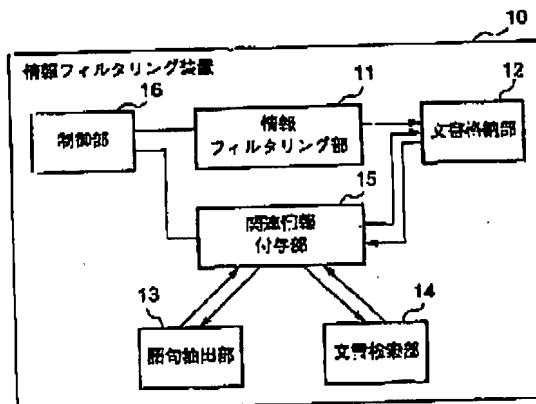
(74) 代理人 弁理士 鈴江 武彦 (外6名)

(54) 【発明の名称】 情報フィルタリング装置および同装置に適用される関連情報提供方法

(57) 【要約】

【課題】 予め登録された検索条件によって選択される情報からその関連情報へのアクセスを簡便にできるようにする情報フィルタリング装置を提供する。

【解決手段】 関連情報付与部 15 は、情報フィルタリング部 11 により選択され文書格納部 12 に格納された文書を語句抽出部 13 に引き渡し、この文書中から特定のカテゴリの語句を抽出させる。次に、関連情報付与部 15 は、語句抽出部 13 が抽出した語句を文書検索部 14 に引き渡し、この抽出された語句に基づいた関連文書の検索を行なわせる。そして、関連情報付与部 15 は、この文書検索部 14 が検索した関連文書の格納位置を含む関連情報を文書格納部 12 に格納された選択文書に付加する。これにより、フィルタリング結果として提供される情報に対する関連情報の検索作業に費やされるユーザの負担を大幅に軽減することを可能とする。



(2)

特開平11-53392

【特許請求の範囲】

【請求項1】 複数の文書の中から予め登録された検索条件に合致する文書を選択する情報フィルタリング装置において、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出する語句抽出手段と、
前記語句抽出手段が抽出した語句に基づいて前記関連文書を検索する文書検索手段と、
前記文書検索手段が検索した関連文書の格納位置を含む関連情報を前記選択文書に付与する関連情報付与手段とを具備してなることを特徴とする情報フィルタリング装置。

【請求項2】 前記文書検索手段は、前記語句抽出手段が抽出した語句を鍵として文書を検索するとともに、その検索された文書に関係づけられた文書を予め定められた範囲内で検索し、これらの検索された文書すべての中から前記選択文書に適した関連文書を選択する手段を有することを特徴とする請求項1記載の情報フィルタリング装置。

【請求項3】 選択された文書自体またはその文書の中から前記語句抽出手段が抽出した語句を履歴情報として記憶する文書履歴記憶手段をさらに具備し、
前記文書検索手段は、前記語句抽出手段が抽出した語句と類似する語句を含んだ履歴情報を前記文書履歴記憶手段が記憶する履歴情報の中から検索し、その検索された履歴情報を用いて関連文書を検索する手段を具備してなることを特徴とする請求項1または2記載の情報フィルタリング装置。

【請求項4】 複数の文書の中から予め登録された検索条件に合致する文書を選択する情報フィルタリング装置に適用される関連情報提供方法において、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出するステップと、
前記抽出された語句に基づいて前記関連文書を検索するステップと、
前記検索された関連文書の格納位置を含む関連情報を前記選択文書に付与するステップとからなることを特徴とする関連情報提供方法。

【請求項5】 複数の文書の中から予め登録された検索条件に合致する文書を選択する情報フィルタリング装置に適用される関連情報提供方法において、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出するステップと、
前記抽出された語句を鍵として文書を検索するとともに、その検索された文書に関係づけられた文書を予め定められた範囲内で検索するステップと、
前記検索された文書すべての中から前記選択文書に適した関連文書を選択するステップと、
前記選択された関連文書の格納位置を含む関連情報を前記選択文書に付与するステップとからなることを特徴とする

する関連情報提供方法。

【請求項6】 複数の文書の中から予め登録された検索条件に合致する文書を選択するとともに、過去に選択された文書自体またはその文書の中から抽出された語句を履歴情報として記憶する情報フィルタリング装置に適用される関連情報提供方法において、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出するステップと、
前記抽出された語句と類似する語句を含んだ前記履歴情報を検索し、その検索された履歴情報を用いて関連文書を検索するステップと、
前記検索された関連文書の格納位置を含む関連情報を前記選択文書に付与するステップとからなることを特徴とする関連情報提供方法。

【請求項7】 複数の文書の中から予め登録された検索条件に合致する文書を選択するコンピュータ上で関連情報を提供するためのプログラムであって、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出し、
前記抽出された語句に基づいて前記関連文書を検索し、
前記検索された関連文書の格納位置を含む関連情報を前記選択文書に付与するように前記コンピュータを動作させるプログラムを記録したコンピュータ読取り可能な記録媒体。

【請求項8】 複数の文書の中から予め登録された検索条件に合致する文書を選択するコンピュータ上で関連情報を提供するためのプログラムであって、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出し、
前記抽出された語句を鍵として文書を検索するとともに、その検索された文書に関係づけられた文書を予め定められた範囲内で検索し、
前記検索された文書すべての中から前記選択文書に適した関連文書を選択し、
前記選択された関連文書の格納位置を含む関連情報を前記選択文書に付与するように前記コンピュータを動作させるプログラムを記録したコンピュータ読取り可能な記録媒体。

【請求項9】 複数の文書の中から予め登録された検索条件に合致する文書を選択するとともに、過去に選択された文書自体またはその文書の中から抽出された語句を履歴情報として記憶するコンピュータ上で関連情報を提供するためのプログラムであって、
選択された文書に関連する文書を検索するための語句を前記選択文書の中から抽出し、
前記抽出された語句と類似する語句を含んだ前記履歴情報を検索し、その検索された履歴情報を用いて関連文書を検索し、
前記検索された関連文書の格納位置を含む関連情報を前記選択文書に付与するように前記コンピュータを動作さ

(3)

特開平11-53392

せるプログラムを記録したコンピュータ読取り可能な記録媒体。

【請求項10】 特定または不特定の利用者に対して予め定められた範疇の情報を提供する情報提供装置において、
提供する文書に関連する文書を検索するための語句を前記提供文書中から抽出する語句抽出手段と、
前記語句抽出手段が抽出した語句に基づいて前記関連文書を検索する文書検索手段と、
前記文書検索手段が検索した関連文書の格納位置を含む関連情報を前記提供文書に付与する関連情報付与手段とを具備してなることを特徴とする情報提供装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、たとえば新たに発生した情報の中からユーザが所望する情報のみを選択して提供する情報フィルタリング装置および同装置に適用される関連情報提供方法に係り、特に選択した情報に関連する情報を適切に検索して付与する情報フィルタリング装置および関連情報提供方法に関する。

【0002】

【従来の技術】近年のインターネットの普及に伴ない、誰もがこのインターネットに接続された世界中に点在する計算機に格納された情報を簡単にアクセスできるようになってきている。特に、WWW (World Wide Web) では、HTTP (Hyper Text Transfer Protocol) を用いることにより、利用者が世界中の情報をGUI (Graphical User Interface) ベースのブラウザによって簡単にアクセスできる仕組みを提供している。このWWWでは、ある計算機上でhttpdと呼ばれるソフトウェアを用いる。このソフトウェアは、自身が動作する計算機のデータベースに格納されているHTML (Hyper Text Markup Language) で記述したハイパーテキストファイルを、他の計算機からの要求に応じて転送するものである。したがって、インターネットに接続されている計算機は、転送してもらいたいハイパーテキストファイルを格納している計算機上で動作しているhttpdに対し、ハイパーテキストファイルのアドレスを指定することによって指定したファイルを読み取ることができる。

【0003】HTMLの記述では、ハイパーテキストファイル内のリンク情報が、前記アドレスにより記述されるので、HTTPのプロトコルにしたがったブラウザは、各httpd支配下のハイパーテキストファイルを表示できることになる。ブラウザは、音声、静止画および動画などのマルチメディアデータを含むハイパーテキストを表示することができる。

【0004】このWWWの仕組みにより、ユーザは、簡単にインターネット上の情報にアクセスできるように

り、多くの個人や企業が、Webページまたはホームページと呼ばれるハイパーテキストファイルを公開するようになってきている。

【0005】一方、インターネットの普及とともに文書の電子化も進んでおり、新聞や雑誌といった出版物も電子的に提供する（つまり電子メールやWWWによって提供される）ことが始まりつつある。従来の紙ベースで行われていた出版が電子化されることにより、今後は書店などにわざわざ足を運ばなくても、家庭にいながら種々の情報を入手できるようになってくると考えられる。しかしながら、このような情報アクセスの利便性の増大は、逆に情報洪水ともいえる状況を作り出しつつある。すなわち、アクセスできる情報があまりにも膨大になる結果、不要な情報の中に肝心な必要とする情報が埋もれてしまい、見失ってしまうといった現象が出てくる。このため、膨大な情報からユーザの要求・興味にあったものを選出してユーザに提供する情報フィルタリング装置の需要が高まりつつある。

【0006】

【発明が解決しようとする課題】しかしながら、従来の情報フィルタリング装置においては、予め定められた情報源（たとえば、新聞社で発行する新聞記事）から情報を選択するだけであった。たとえば、新聞社が発行する新聞記事を対象とした情報フィルタリング装置では、ユーザごとに設定された検索条件にしたがって関連する記事を選択するのみである。このため、仮にフィルタリング結果の記事内の内容に関してより詳しいことを知りたいとユーザが感じても、その記事から直接情報を得ることができないという問題があった。

【0007】たとえば、わからない単語が出てきた場合は、その単語の意味を辞書などで調べる必要がある。このような場合、ユーザは辞書を引けば良いだけであるので、たいした手間がかかるとはいえない。ところが、多くの場合、単語の意味内容を調べるといった単純なものではなく、関連する情報を収集したいといったことが多い。したがって、たとえばWWW上で公開されている検索サーバを使って、知りたい内容をさらに検索するといったようなことを行なう必要があった。あるいは、記事が新製品の発売などの記事であった場合には、その記事内容を発表した企業のホームページをWWW上で閲覧し、そこに格納されているプレスリリースの原稿や関連情報などを探して必要な情報を得る必要があった。

【0008】単語の意味を調べるといった単純な内容であれば辞書で十分であるが、関連情報を調べるといったときに、そのような情報は辞書や百科事典といった出版物に頼ってはいない新鮮な情報を得ることができない。現在、WWWが普及してきた状況にあつては、新鮮な情報を得ようとするれば、WWWを利用するのが最適な方法であるといえる。

【0009】しかしながら、現状のネットワーク基盤に

(4)

特開平11-53392

あつては、WWWでWebページにアクセスするには多くの通信時間がかかるため、ユーザが所望の情報を得ようとすると、膨大な手間と時間が必要である。

【0010】この発明はこのような実情に鑑みてなされたものであり、記事中の詳細情報や関連情報などを掲載しているWebページを事前に検索しておき、そのようなページとのリンク付けを予め行なっておくことにより、ユーザの手間を格段に減らすことを可能とする情報フィルタリング装置および同装置に適用される関連情報提供方法を提供することを目的とする。

【0011】

【課題を解決するための手段】この発明は、前述した目的を達成するために、選択された文書を解析して特定のカテゴリの語句を抽出し、この抽出された語句に基づいて関連文書を検索するとともに、この検索された関連文書の格納位置を含む関連情報を選択文書に付与するようにしたものである。

【0012】この発明においては、たとえばあるWebページをユーザに提供する場合、その関連情報を載せたWebページを適切に検索してそのアドレスを埋め込むことにより、ユーザは簡単な操作のみでその関連情報を載せたWebページを表示できるようになる。このため、検索サーバによってWebページを検索したり、Webページを表示させながら関連する情報の有無を調べるといったユーザの負担を大幅に軽減させることが可能となる。

【0013】

【発明の実施の形態】以下、図面を参照してこの発明の実施形態を説明する。

(第1実施形態) まず、この発明の第1実施形態を説明する。図1に、この第1実施形態に係る情報フィルタリング装置の機器構成を示す。図1に示すように、この第1実施形態の情報フィルタリング装置は、フィルタリング部11、情報フィルタリング部11で選択された文書を格納する文書格納部12、文書格納部12に格納された文書から関連情報を検索する手掛かりとなる語句を抽出する語句抽出部13、抽出された語句に基づいて文書を検索する文書検索部14、検索された文書が格納されているアドレスを処理対象の文書に付与する関連情報付与部15、および情報フィルタリング部11を起動した後、関連情報付与部15を起動する制御を行なう制御部16からなる。また、この第1実施形態の情報フィルタリング装置10では、情報フィルタリング部11、語

句抽出部13、文書検索部14、関連情報付与部15および制御部16は、記憶装置2に記憶されてCPU1によって実行制御されるプログラムとして構成され、また、文書格納部12は、記憶装置2上に構築されている。

【0015】なお、情報フィルタリング部11における処理内容は、この発明の主旨ではない。たとえば、特願平7-212940号などで開示されている方法の一部を流用することにより実現することができる。つまり、たとえば新聞記事に対する情報フィルタリングでは、新聞社などから電子的に配信される新聞記事を対象に、予め登録されている検索条件との類似度計算を行なう。そして、この類似度計算で得られる類似度値が予め定められた閾値を越える記事を選択記事として文書格納部に格納する。また、この特願平7-212940号で開示されている情報フィルタリング装置では、選択された記事を類似度の順にソーティングしてユーザに提供する。一方、これらに対してこの第1実施形態の情報フィルタリング装置10では、文書格納部12に格納された記事についての関連情報を自動的に検索し付与する点を特徴としている。

【0016】ここで、文書格納部12の格納形式を図3(a)に、格納された記事の一例を図3(b)に、図3(b)に示した記事に対応するHTMLを図3(c)にそれぞれ示す。

【0017】次に、関連情報付与部15の処理の流れを図4に図示する。関連情報付与部15は、文書格納部12に格納されている記事を対象に、語句抽出部13を起動し検索の手掛かりとすべき語句を抽出する(ステップA1)。次に、関連情報付与部15は、抽出した語句を文書検索部14に送り、関連ページを検索する(ステップA2)。そして、関連情報付与部15は、関連ページのアドレスを対象となっている記事に付与し処理を終える(ステップA3)。

【0018】記事から単純に語句を抽出し、それを使って検索を行なうと、ユーザにとって必要でないページまでも検索してしまうという問題がある。そこで、精度よく検索を行なうことを目的とした関連情報付与部15の処理の流れを図5に示す(図4の処理の流れを新聞記事を対象に関連するWebページを検索するものとして詳細化したものである)。

【0019】まず始めに、関連情報付与部15は、記事から主体者となる語句を抽出する(図3(b)の記事では「日本XYZ」を抽出する)(ステップB1)。次に、関連情報付与部15は、この取り出した主体者名からその主体者が公開しているホームページを求める(ステップB2)。具体的には、たとえば図6に示したような形式で、会社名とその会社が公開しているホームページのURLとの対応テーブルを用意しておく。そして、この会社名と抽出した主体者名とを照合することによつ

(5)

特開平11-53392

て、その主体者が公開しているホームページのアドレスを求めることは容易に実現できる。

【0020】次に、関連情報付与部15は、この求めたURLを起点として関連ページを検索し(ステップB3)、関連ページのアドレスを記事に埋め込む(ステップB4)。ここで、図3(b)の例に対して、検索された関連ページのURLが仮にhttp://www. xyz. co. jp/release/970324. htmlであると仮定すると、関連ページへのアドレスは、HTMLの記述法を用いて、図7の22a~22bのように埋め込むことができる(ホームページのアドレスは、図7の21a~21bのように埋め込まれる)。

【0021】次に、語句抽出部13の処理の流れ(図5のステップB1に対応)を図8に示す。この第1実施形態の語句抽出部13では、記事における主体的な位置を占める会社名を抽出するために、表現パターンを用いて抽出する。表現パターンとしては、図9に示すような、新聞記事において会社名を表現する表現パターンと会社名に対応する部分を表現した規則とをテーブルとして記憶しておく。図9において、「^」は、文の始めを意味しており、「*A」や「*B」は、任意の文字列を意味している。会社名フィールドの「*A」は、表現パターンフィールドのパターンのうち、会社名に対応する部分を表現している。語句抽出部13では、このような表現パターンと記事とを照合し、照合に成功した場合に、会社名に対応する文字列を取り出す(ステップC1~ステップC2)。

【0022】たとえば、図3(b)に示した記事は、図9に示す表現パターンの1行目と照合する。つまり、変数「*A」には「日本XYZ」が、変数「*B」には「東京都千代田区〇〇町××、社長□□氏」が照合することになる。したがって、会社名に対応する「日本XYZ」が出力として得られる。

【0023】また、表現パターンとして図10に示すような商品名を抽出する規則を用意しておけば、商品名を抽出することができる。ここで述べた処理では、単純な文字列照合による方法を述べたが、たとえば形態素解析を行なうことにより、語の区切りを求めて精度を向上させることも可能である。

【0024】Webのページは、ハイパーテキスト構造になっており、図11(a)に模式的に示したように、各ページが関係づけられている。そして、HTMLの記述法では、図11(b)に示すように、というタグでリンクづけするページのURLを記述することで関係づけが行なわれている。企業や個人が公開するページは、Webブラウザによって大きな話題から詳細な話題へとユーザがリンクをポインティングしながら所望のページへ辿っていくことを想定しており、図11(a)に示すような階層的な構成とする場合が多い(リンクの関係がループとなる場合も存在す

る)。

【0025】ここで、文書検索部14の処理の流れ(図5のステップB3に対応)を図12に、再帰的に処理するサブルーチンを図13に示す。文書検索部14は、語句抽出部13で得られたURLを起点として、まず、そのURLのページを読み込む(ステップD1)。次に、文書検索部14は、その読み込んだページと関連ページ検索の元となっている記事との類似度を求め、類似度が一定の値を越える場合に関連ページの候補とする(ステップD2~ステップD4)。

【0026】ここで、文書検索部14は、そのページ内にリンクが含まれるか否かをチェックし(ステップD5)、含まれる場合は(ステップD5のYES)、それらの各リンクについてサブルーチン呼び出す(ステップD6~ステップD7)。このサブルーチンは、図13に示すように再帰的な関数である。そして、文書検索部14は、最終的に得られた候補の中から最大の類似度を持つページを関連ページとして選択する(ステップD8)。

【0027】図13に示すサブルーチンでは、まず始めに、リンクのループにより同じページを何度も処理することを避けるため、入力されたURLがすでに処理したものか否かをチェックする(ステップE1)。また、他のWebサーバへのリンクをたどることは避けることにより、検索の範囲が発散することを避ける(ステップE2)。さらに、階層をたどる深さがあまり深くならないように、既定値を越えたか否かをチェックする(ステップE3)。そして、これまで処理していないURLであり、他のサイトへのリンクでもなく、かつ深さが既定値以内のURLであれば、入力されたURLのページを読み込むとともに(ステップE4)、入力記事との間の類似度を求める(ステップE5)。そして、類似度が所定の値を越える場合に(ステップE6のYES)、関連ページの候補として格納する(ステップE7)。さらに、そのページがリンクを含む場合は(ステップE8のYES)、それらの各リンクに対して、自分自身を再帰的に呼び出す(ステップE9~ステップE10)。

【0028】なお、入力記事と読み込んだページとの間の類似度の算出法は、たとえば特開平7-212940号などで開示されている方法を用いることができる。また、この第1実施形態では、記事全体との類似度を求めることにより、関連ページを検索する方法を取ったが、これを商品名のように特定の語句により検索するように変形することは容易である。

【0029】(第2実施形態)次に、この発明の第2実施形態を説明する。図14は、この第2実施形態の情報フィルタリング装置の機能ブロック図である。

【0030】この第2実施形態の情報フィルタリング装置は、第1実施形態と比較して、文書履歴記憶部17を有すること、および関連情報付与部15の処理の流れが

(6)

特開平11-53392

異なる。以下、これらの異なる部分について説明する。

【0031】関連情報付与部15の処理の流れを図15に図示する。この処理では、関連ページの検索にあたり、文書履歴記憶部17に格納されているこれまでの入力文書から類似度の強い文書を検索し、その結果に基づいて、実際の検索を行なうものである。これは、たとえば提供しようとする記事に先駆けて提供された記事の中から関連する記事を検索し、その検索された記事の主体者となっているホームページを起点に関連ページを得ることを目的にしている。このように2段階で行なうことによって、提供しようとする記事の関連情報を載せたWebページへのリンクが効率的に得られることになる。

【0032】すなわち、関連情報付与部15は、文書格納部12に格納されている記事を対象に、語句抽出部13を起動し検索の手掛かりとすべき語句を抽出する（ステップD1）。次に、関連情報付与部15は、この抽出した語句と関連する項目が文書履歴記憶部17に存在するかを検索する（ステップF2）。文書履歴記憶部17に関連する項目が存在した場合（ステップF2のYES）、関連情報付与部15は、文書履歴記憶部17から取り出した項目を文書検索部14に送り、関連ページを検索する（ステップF3）。そして、関連情報付与部15は、関連ページのアドレスを対象となっている記事に付与する（ステップF4）。また、関連情報付与部15は、語句抽出部13が抽出した語句を文書履歴記憶部17に格納しておく（ステップF5）。

【0033】文書履歴記憶部17の記憶形式の例を図16に示す。ここで、「その他」の部分については、文書中で用いられた会社名や商品名以外の名詞を設定しておくものとする（抽出した語句ではなく記事全体を格納しておいても構わない）。

【0034】この図16に示したような語句を文書履歴として格納しておく、図3(b)の記事との類似度を求めることにより1の項目が選択できる。関連情報付与部15は、この項目を文書検索部14に送り、関連ページを検索する。そして、その検索された関連ページのアドレスを、たとえば図17に示す領域23のように、提供される記事に埋め込む。これにより、より効率的に関連情報を提供することが可能となる。

【0035】なお、関連ページへの自動的なリンクづけは、第1および第2実施形態の情報フィルタリング装置のように、文書を選択する装置のみに限らない。一定の文書を提供する情報提供装置の場合に関しても、語句の抽出および文書検索を行なうことにより、関連情報を付与することは有効である。

【0036】また、この発明の手法をサーバ側で実現し、関連情報を付与した文書を電子メールで送付することも可能であるし、ユーザ側のクライアントがHTTPのプロトコルにしたがって読み出すように実現することも可能である。また、単純なシステムとして実現した場

合には、個人のパーソナルコンピュータやワークステーションなどで動作し、文書格納部に格納されている（関連情報が付与された）文書を、ハイパーテキストを表示するためのブラウザで読み出させることでも実現可能である。

【0037】さらに、この発明の手法は、ソフトウェアとして実現可能であるため、CD-ROMやフロッピーディスクなどといった記録媒体によって頒布することが可能である。また、磁気ディスクなどに格納しておき、ネットワークで取り寄せる（ダウンロード）ような形式で頒布することも可能である。

【0038】

【発明の効果】以上詳述したように、この発明によれば、複数の文書の中から予め登録された検索条件に合致する文書を選択して提供する場合、この提供しようとする文書の関連情報を格納したWebページのアドレスを自動的に付与することにより、ユーザは簡単にそのページをブラウザによって表示できるようになる。このため、検索サーバによってページを検索したり、ページを表示させながら関連する情報の有無を調べるといったユーザの負担を大幅に軽減することが可能となる。

【図面の簡単な説明】

【図1】この発明の第1実施形態に係る情報フィルタリング装置の機器構成を示す図。

【図2】同第1実施形態の情報フィルタリング装置の機能ブロック図。

【図3】同第1実施形態の文書格納部の格納内容の一例を示す図。

【図4】同第1実施形態の関連情報付与部の処理の流れを示す図。

【図5】同第1実施形態の精度よく検索を行なうことを目的とした関連情報付与部の処理の流れを示す図。

【図6】同第1実施形態の会社名とホームページのURLとの対応テーブルの一例を示す図。

【図7】同第1実施形態の関連情報を埋め込む例を示す図。

【図8】同第1実施形態の語句抽出部の処理の流れを示す図。

【図9】同第1実施形態の語句を抽出するための表現パターンを例示する図。

【図10】同第1実施形態の語句を抽出するための表現パターンを例示する図。

【図11】Webのページのハイパーテキスト構造を説明するための図。

【図12】同第1実施形態の文書検索部の処理の流れを示す図。

【図13】同第1実施形態の文書検索部の処理の流れ（サブルーチン）を示す図。

【図14】この発明の第2実施形態に係る情報フィルタリング装置の機能ブロック図。

(7)

特開平11-53392

【図15】同第2実施形態の関連情報付与部の処理の流れを示す図。

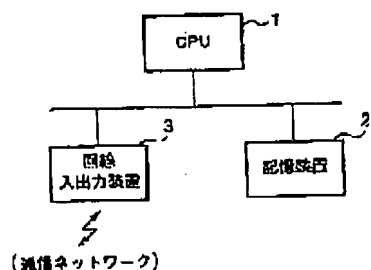
【図16】同第2実施形態の文書履歴記憶部の記憶形式の例を示す図。

【図17】同第2実施形態の関連情報を埋め込む例を示す図。

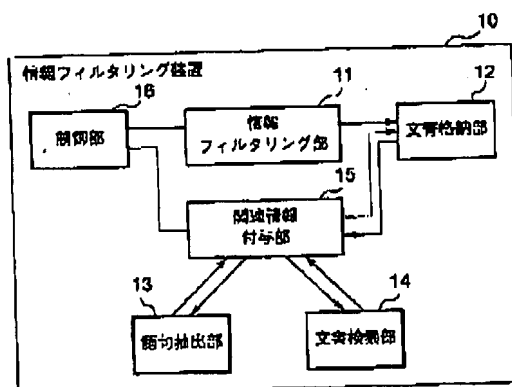
【符号の説明】

1…CPU、2…記憶装置、3…回線入出力装置、10…情報フィルタリング装置、11…情報フィルタリング部、12…文書格納部、13…語句抽出部、14…文書検索部、15…関連情報付与部、16…制御部、17…文書履歴記憶部。

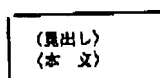
【図1】



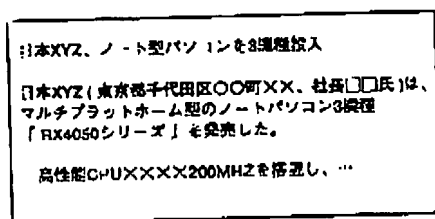
【図2】



【図3】

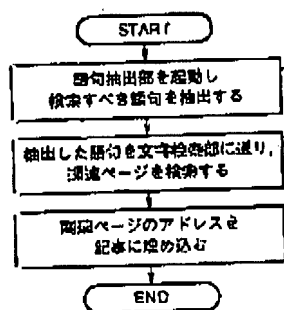


(a)

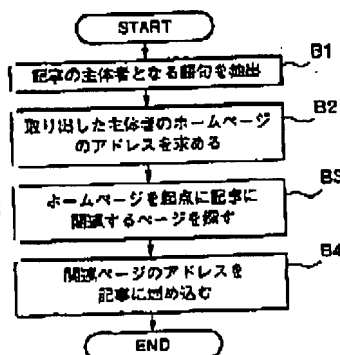


(b)

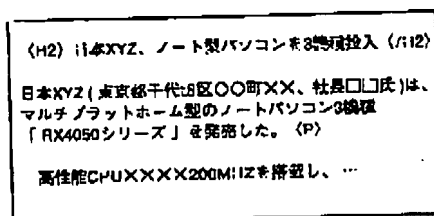
【図4】



【図5】



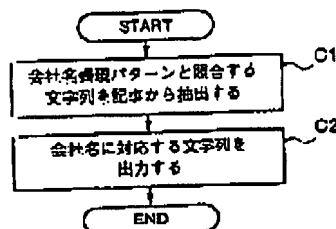
【図6】



(c)

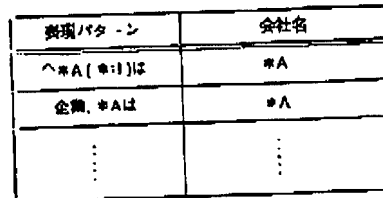
会社名	URL
AAA	http://www.aaa.co.jp
BBB	http://www.bbb.co.jp
⋮	⋮
日本XYZ	http://www.xyz.co.jp
⋮	⋮

【図8】

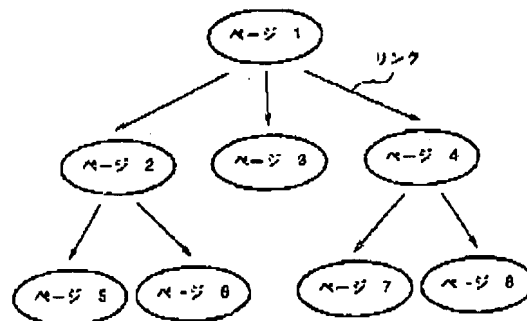


特開平 1 1 - 5 3 3 9 2

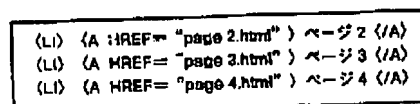
【圖9】



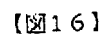
【 1 1 】



(८)



(b)



1	会社名: 日本〇××, 商品名: PX3000 その他: マルチプラットフォーム、ノートパソコン、発売、...
2	会社名: ×××, 商品名: □□□ その他: 液晶、カメラ、投入、...
!	!